Erstveröffentlichung in ZfBB 54 (2007) 4-5, S.268-272. S. 268
Die Seitennummerierung der Originalveröffentlichung ist kenntlich gemacht, so dass der Beitrag damit entsprechend zitierbar ist.

Open Access is a Choice

Jan Velterop
Springer Science+Business Media, Springer UK
Director of Open Access

Die formale Veröffentlichung von Forschungsergebnissen mit einem Peer-Review-Verfahren wird in der Wissenschaft als notwendig erachtet. Zugleich wird der freie Zugang zu diesen Ergebnissen zunehmend als notwendig erachtet. Manchmal werden diese Bedürfnisse als unvereinbar empfunden, sie sind es aber nicht. Wir müssen von der konventionellen Rolle, die Urheber-recht und Abonnements spielen, wegkommen, um die Möglichkeiten des Publizierens als Dienstleistung voll auszuschöpfen und zugleich freien Zugang zu erhalten.

Formal peer-reviewed publication of research results is seen as a necessity in science. But also, open access to these results is increasingly seen as a necessity. The two are sometimes seen as incompatible, but they are not. But we have to get away from the conventional roles of copyright and subscriptions in the system in order to capitalise on the possibilities of publishing as a service, with open access as a result.

Introduction

Information is funny stuff. You can give it to someone, and still keep it. You can sell it to someone and still keep it. It multiplies like the biblical loaves and fish. Not food, but food for thought. It is not used up when it is consumed. It can be simultaneously perishable and non-perishable (though not quantum physics style), for if it has 'newness' it is often very quickly stale, perishable, but its actual information content is not. That is often very carefully archived and kept for future reference. Information cannot, for all these reasons, be regarded and treated in the same way as goods. And though it is often seen as 'intellectual property', it is a very different kind of property when compared to 'real' property.

Because of all the conventions that surround the publication of scientific research information, that is even funnier stuff. You can find pretty much the same news in different newspapers. Global news appears in every newspaper. Not so in science. The convention that a scientific article is published only once makes the information *non-rivalrous* – it makes it unique, not to be found in any other journal than in the one in which it is first published, and if it is obtained from anywhere else, it will be no more than a copy, carrying the original journal's imprimatur. This makes sense, because scientific articles are accumulative, they accrete knowledge, connect bits and pieces, and they do that via references and citations. If the same article could be referred to as having been published in more than one journal, it would be a right mess and the coherence of the network of scientific knowledge would be jeopardised. But it also means that journals are *monopoloid* in nature. They don't compete for readers in the same way as newspapers or magazines do. Instead, they compete for authors who submit their articles for publication.

So the economics are funny, too. Scientific journals are usually sold on subscription, but they don't compete for subscriptions, where their income comes from; they compete for authors, who don't pay. They can't compete for subscriptions, because the readers can't choose. Only authors can choose. Suppose there are two journals, in the same field, in which important research is being published. Authors can choose to submit to one of them. Not only *can* they choose, in fact, they *must* choose. They could only decide to submit the same article to another journal after it has been rejected by the first. But how could a researcher in that field possibly decide to read one journal and not the other? If he needs the one, he will need the other. Researchers need comprehensiveness. They can't afford the risk of missing important articles just because they, or their libraries, don't have a subscription to all the journals.

In reality, of course, true comprehensiveness is just a dream. There are so many articles being published in so many different journals, that few, if any, libraries in the world have a subscription to all of them or indeed could afford to. That's why a compromise has developed over the last century. Researchers can get a reasonable idea of what's been published in their field, because abstracts of most papers have usually been made freely available. A whole industry of abstracting and indexing services has grown up in order to offer this much needed comprehensive overview, across journals and publishers. With few exceptions, abstracts have for a long time already been 'open access' and the internet has made that access only easier.

But abstracts are only abstracts. Abstracts give a researcher not much more than the information that there is an interesting article he ought to read, and what it is about, but if he or his library has no subscription, he still can't see the full article and all the detailed information in it. Wouldn't it be nice if he could have full-text open access to any article of interest? That would make sense, wouldn't it? What could be done to bring that about?

Open Access

Full-text open access *per se* is easy nowadays. Anybody can publish whatever they like on the internet and

there are plenty of opportunities to do so at no cost, or virtually no cost other than that of a basic connection to the internet. Published like that, in what used to be called *cyberspace*, any information can freely flow to just about any recipient in the world who is connected to the web. But especially within peer-groups communication has become exceedingly easy. Numerous blogs, email discussion lists, wikis, and similar forums exist that facilitate it and are being used for it. Google blog search for instance (http://blogsearch.google.com) is likely to find at least one active blog in virtually any discipline, and many blogs in most fields. And there are providers of blogs specifically aimed at science, such as Scienceblogs (http://scienceblogs.com/). Authors of scientific articles could easily use these opportunities. The whole scientific discourse could comprehensively take place on the internet, without access barriers, and indeed a large amount of information is already shared in the scientific community in that way.

Formal publishing

Open access publishing – structural open access to the formally published article, from the day of publication in peer-reviewed journals – has been presented as a solution to the research communication problem. However, in the light of what has been said above, it hardly is a solution for current scientific communication and the exchange of ideas, because at least in many fields, such communication does not seem to have much of a problem for which it needs a solution.

So why do we need open access publishing? What then is the problem to which open access publishing *is* the solution?

Whilst the free availability of scientific research results is widely desired, the prevailing thought in the academic culture is that these results can only be trusted as authoritative if they are formally published in peer-reviewed scientific journals. The world of research relies on trust. Peer review, though not perfect, is a reasonable way to ensure the veracity, the scientific soundness, of the building blocks used in subsequent research, and peer-reviewed journals provide the 'trust-labels'. Whilst the communication itself may already have taken place before the research article in question was published, or even before it was submitted, the availability of a 'label of trustworthiness' is of great value. Open access publishing ensures not only universal online access to research articles, but to research articles with such formal trust labels.

Put starkly, formal science publishing is not all about communicating research results; it is about communicating how trustworthy are research results. Not strictly about content, in other words, but about what could be called 'meta-content'. This meta- content gives, via the mechanisms of peer-review, information about, for instance, whether the conclusions are warranted by the data presented;

whether the method applied proper and fit for the purpose; by whom – by which sub-community – the article was judged to be trustworthy and publishable. This meta- content, the answers to these questions, provides proof that the article was accepted by the community as up to their usual standards of scientific soundness, of plausibility given the current state of knowledge (though not necessarily of truth; that's not how science works). Formal peer-reviewed journals provide this meta-information, if not explicitly, then at least implicitly by having accepted the article for publication.

Economics

However, unlike much of the informal communication, the service of formal science publishing costs money. Publication in a peer-reviewed journal takes a fair amount of organisation and work. And if those tasks are to be taken on professionally, the people who do the organisation and work that is involved need to be paid, since for them it is a day job. On top of that, there is a need for all manner of technical and infrastructural provision that needs to be in place to get the job done, and done well and in time. This is often underestimated, and, to be fair, it is mostly invisible. But like scars after cosmetic surgery, it is invisible precisely *because* it's done so well and so professionally.

Publishers have traditionally covered the costs they incur in the process of formal peer-reviewed publication by selling subscriptions to their journals. Later on, when the journals were also published electronically on line, these subscriptions became licences, but that is a mere legal difference and not a principle one, at least not unless you are a lawyer. The basic idea remained, and that is that the reader – or rather, the university librarian on behalf of the reader – pays the bill.

Then came open access. The very principle of open access entails that subscriptions, access licences, or any restrictions to free access are anathema. That being so, they cannot be used in any way to generate the income needed to defray the cost of publishing. For the printed version of journals one can still maintain a subscription system, to pay for the real marginal cost of printing and distribution, but in the case of open access articles such a subscription doesn't pay for the content – after all, the content is freely available – but only for the *convenience* of having the articles neatly

printed and bound in journal issues or volumes. The market for just this *convenience* is likely to be smaller than the market for the content itself, or than the market for the content in combination with the convenience, and so the income potential for a journal from subscriptions is also smaller. Much smaller, in all likelihood, and in most cases too small to cover the costs a publisher needs to incur in order to keep a journal going.

For some very small journals, the costs could be – and are – subsumed in academic departments of universities and the necessary work done by volunteers. In those cases, online open access could be gratis. In cases where a more professional approach is needed, or deemed appropriate, and therefore a viable income stream is required, new journals were started on an open access model sustained by a charge levied on the author whose articles were to be published in the journal (though this charge is rarely paid by the author, but, instead, by the funding agency that supports the research, or by the author's institution).

This works well for new journals. Not, however, for existing journals with a loyal authorship. If an established journal suddenly were to start charging all its authors a fee for publishing, it would run the risk of sending some of its authors away to other journals that don't (yet) charge such fees. No publisher can be expected to do that. However, not all authors would run away, because some would value open access to their article enough to be willing to pay the fee. Or they would be encouraged by their funders to pay the fee and so obtain full open access.

The Hybrid model

The first to realise the consequences of that was Thomas J. Walker, Editor-in-Chief of the *Florida Entomologist*. He first proposed that authors be given the choice[1] and then he actually implemented the authors' option at the *Florida Entomologist*. Initially the way it was presented was to 'sell' electronic 'reprints' to authors, which they then could post anywhere on the internet, in effect allowing

them to make the articles freely accessible to anybody in the world. This was before the term 'open access' was coined, after all.

This 'Walker-model', now often described as the 'hybrid' model, is the essence of Springer's Open Choice[2]. Springer was the first large publisher to apply this model to the entire range of its journals. In this model, the publisher doesn't impose anything on the author, but, after the article is accepted for publication following a process of peer review, the author is given the option to publish his or her article with open access. If authors choose that option, they pay a fee of 3000 $ or equivalent, and their respective articles will be fully open. They will be published with an open access licence substantially identical to the Creative Commons Attribution License[3] which means that the official, published version of the articles can be freely distributed by anyone, anywhere, in print or on line, as long as the authors and original source are properly acknowledged. This includes posting on the author's or an institutional web site, or any other web site for that matter, and photocopying. The Springer Open Choice programme applies to all of Springer's own journals and to the majority of the journals copublished with others.

The cost issue

Some open access advocates have an agenda that has more to do with reducing the cost of science publishing than with access. They argue that the functions of certification, organising peer review, and that of dissemination, should be split. I can agree that they certainly could be split. The mistake that's often made is to assume that the bulk of the cost of publishing is associated with dissemination. In fact, most of the cost of formal publishing is to do with certification, and everything that needs to be done for the article to be fit for dissemination. This is known as the 'firstcopy cost'. Costs associated with multiplication and distribution, are marginal costs and variable with the number of copies made and disseminated in print.

Even when they get it right, and realise that the first copy costs are the bulk of the per-article costs, there is a problem with a cost-based approach. The so-called 'serials crisis' is, after all, the result of just such a cost-based approach.

To illustrate this: suppose a journal would operate purely on a cost-recovery basis. It has n subscriptions, publishes m articles per year, and charges an annual subscription price of x. Its total costs (Ct) would be nx; its per-article costs nx/m. If Ct remains constant in real terms, and n, too, this system could be quite stable. However, when subscriptions fall, i. e. n decreases, the price, x, would have to go up to cover the costs. If n becomes $0.8n$, then x needs to be $1.25x$. In other words, in a cost-based system, a 20 % decrease in numbers of subscriptions causes a 25 % increase in price. We have seen decades of subscription attrition, and a large portion of the price increases seen over that period can be ascribed to this mechanism. And because price increases often are a cause of cancellations, the conditions for a vicious cycle are set. The serials crisis, in a nutshell.

Information is easily shared. A library is an institution devoted to sharing information within an en-

tire institution. Sharing between libraries of different institutions is a natural development. Consortia have been formed. Inter-library loan – a rudimentary term left over from the print era, when a loan was indeed a loan – has evolved to become interlibrary document delivery or interlibrary document sharing. There is very little reason to expect further subscription erosion to stop. The result is that, in order to sustain journals, subscription prices are more likely to go up than down.

The quality issue

The quality of open access articles is sometimes called into doubt. However, that is misguided. The quality of the open access articles published under the Springer Open Choice programme is obviously the same as that of traditionally published articles, since the open access option is only given after successful peer-review, which underlies the decision by the Editor-in-Chief to accept for publication. That decision is entirely on scientific merit, just as is the case for traditionally published articles, and in no way informed by commercial considerations.

The open access articles are not only published on line, but are also included in the printed version of the journal, in order to have a complete record in print, as paper archives are still seen as an important way to preserve scientific knowledge.

Difficulties

Just like the full open access model, in which authors are given no choice but to pay and have open access, the hybrid model, in which authors are given the choice, faces some practical difficulties. The first one is that many universities and research institutions are not well set up to deal with payments for article processing fees. Some of that may be perception more than reality. The argument often heard is that there isn't such a tried and tested mechanism as there is for library budgets, for instance. However, traditional payments for page charges, colour charges, reprints, et cetera, are in reality very similar to the article processing fees needed for open access.

The second difficulty is that the amount of research done at a given institution is not always indicative of the amount they traditionally spend on scientific literature. Institutes focussed on research have different needs in that regard to universities focussed on teaching. Similarly, research establishments with a narrow focus have different needs to those with a multi- or cross-disciplinary remit, the latter usually needing a much wider range of journals. Research-intensive institutes typically publish far more articles relative to the ones they read than teaching-intensive universities, where reading prevails over publishing. A shift from payment on behalf of the reader – the subscription model – to one with payment wholly or partially on behalf of the author – the open access and hybrid models – will naturally bring about a shift in financial support for journals from reader-intensive universities, now bearing the brunt of the cost, to research- intensive institutions. Although Academia as a whole would not pay more as a result of open access publishing, those who will find themselves with lower bills understandably are more enthusiastic than those who will face higher bills. This is a considerable hurdle for the success of open access and hybrid models, but the issue is increasingly being taken up by funding agencies, who are taking the sting out of it by defining publication of the results as an integral and necessary part of research itself and therefore the cost of publication as an integral and necessary part of the cost of research. According to this principle, the best-funded institutions pay the most. Given that the money for both subscriptions and article processing charges largely comes from funding agencies in the first place, albeit via a circuitous route, this is a solution to the problem, since the bill for funding agencies remains the same.

There is another hurdle to overcome for the hybrid model. The criticism sometimes heard is that the model makes Academia pay twice: for subscriptions as well as for article processing charges. This is not the case, though. Nobody pays twice for the same articles. The amount the publisher needs for the publication of an article is either covered by subscriptions, or by article processing charges. Never by both. Open access articles that already have been paid for are ignored for the purpose of establishing a journal's subscription price for the following year. So while there may be a difference in phasing, the same material is not charged for more than once. The criticism arises from the fact that this cannot always be made clearly visible. Particularly not if an increase in non-open, traditional articles in a journal materially outstrips the number of open access articles published in it. In that case, the price for the following year may even turn out higher. But if so, that is on account of the extra traditional, non-open articles in it, rather than as a result of its open access articles.

The problem with which both institutional libraries as well as publishers struggle, be they open access, hybrid or traditional publishers, is how to cope with the ever increasing number of qualitatively good research articles submitted for publication and accepted by peer-reviewers. The funding agencies' policy of

regarding the cost of publication as an integral cost of doing research alleviates this problem.

Currently, the authors' uptake of the Open Choice is still relatively small, but it is expected to grow substantially in the years to come.

[1] See *BioScience* 45**,** 171 (1996), www.fcla.edu/FlaEnt/bioscivp.htm [accessed March 9th, 2007].

[2] www.springer.com/openchoice [accessed March 9th, 2007].

[3] http://creativecommons.org/licenses/by-nc/2.5 [accessed March 9th, 2007].